

CONTROL VISUAL PARA EL SEGUIMIENTO DE MÓVILES EN UN PLANO

Jorge Pomares, jpomares@disc.ua.es

Pablo Gil, pgil@disc.ua.es

Fernando Torres, medina@disc.ua.es

Grupo de Automática, Robótica y Visión Artificial (<http://www.disc.ua.es/gava>)

Dpto. Física, Ingeniería de Sistemas y Teoría de la Señal

Escuela Politécnica Superior - Universidad de Alicante

Crtra. San Vicente, s/n, 03080 Alicante

Resumen

En este artículo se analiza un sistema de control visual. En la realimentación se dispone de un sistema de visión artificial compuesto por módulos de bajo y alto nivel atendiendo al tipo de información que procesan. También se analizan distintos tipos de control y se comprueba cual es su resultado en el desempeño global de sistema.

Palabras Clave: Control visual, tracking.

1 INTRODUCCIÓN

El seguimiento de objetos en movimiento es motivo de recientes investigaciones [8] [7] [1] lo que motiva la aparición de nuevos algoritmos orientados al procesamiento de secuencias de imágenes a alta velocidad. El desarrollo de estos algoritmos a su vez implica un mejor desempeño de los sistemas de control visual que los utilizan.

La característica en común de la mayoría de los sistemas de visión clásicos es que son estáticos. Podemos decir que trabajan en bucle abierto, observan el entorno y después realizan el movimiento (lo que se denomina 'look-then-move' [3]).

La exactitud de 'look-then-move' depende de la exactitud del sensor visual y del manipulador. Una alternativa para incrementar la perfección de estos sistemas es usar sistemas visuales realimentados, lo que se suele también denominar visual servoing (término introducido por Hill y Parks [6] en 1979).

Visual servoing consiste en el uso de visión en el más bajo nivel de la jerarquía de control. Para su utilización es necesario que el procesamiento de las imágenes sea rápido y simple. El gran beneficio que se puede obtener utilizando visual servoing es la

exactitud que se consigue con una configuración en bucle cerrado, lo que hace al sistema relativamente insensible a errores de calibración.

El resurgimiento en la actualidad de los sistemas de control visual se debe principalmente al aumento de la velocidad en el procesamiento de imágenes con lo que se ha podido utilizar entornos no estructurados.

Las aplicaciones de este tipo de sistemas son muy amplias, así, encontramos trabajos en la actualidad en los que se describe algunas de ellas como en [5] que usa la alta resolución de un dispositivo óptico, por ejemplo, un microscopio, para realizar el microposicionamiento de las partes implicadas realizando una "servo-visualización". Corke, Roberts y Winstanley [2], describen una aplicación para la industria minera. De nuevo, en [10] encontramos una aplicación de "visual servoing" aplicada a la robótica, en este caso se trata de aplicar a tareas de navegación de un robot el conocimiento proporcionado de la realimentación visual.

En este artículo se va a desarrollar un sistema de control por ordenador para el seguimiento de objetos móviles utilizando visión artificial como sensor. Un móvil será manejado por el usuario y el otro lo controlará el ordenador, realizando un seguimiento según las especificaciones indicadas.

En los siguientes apartados se describen los algoritmos desarrollados y la plataforma de ensayo empleada.

2 DESCRIPCIÓN DEL SISTEMA

2.1 ESQUEMA FÍSICO

En la figura 1 se han representado los distintos componentes implicados en el sistema.

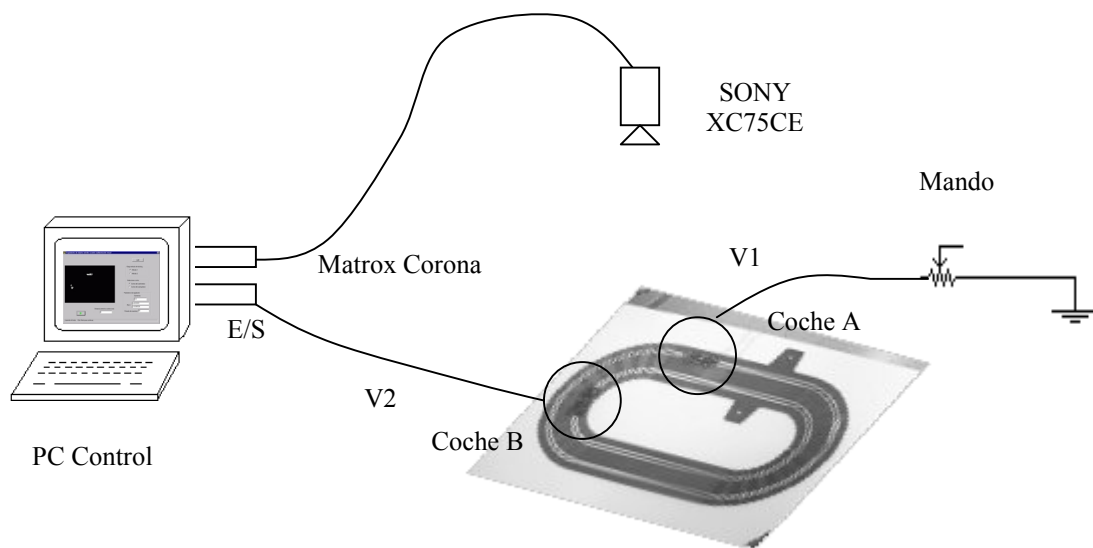


Figura 1: Arquitectura global del sistema

El sistema está compuesto por dos pistas independientes sobre las que circulan dos coches eléctricos. El coche A, que denominaremos perseguido, es manejado por un usuario y es la referencia permanente para el coche B, que es controlado de forma que se mantenga a una distancia fija del perseguido. A este último coche lo denominaremos perseguidor.

Como plataforma física se ha empleado un scalextric modelo Car Racing F1 del fabricante Carrera Century. El coche perseguido es manejado por una fuente de tensión externa (V1). El coche perseguidor es controlado por el PC de Control de manera que el seguimiento se ajuste a las especificaciones impuestas. El ordenador utilizado tiene un procesador Pentium 200 MHz con 32 Mbytes de RAM y S.O. windows NT workstation.

El sistema de visión está compuesto por una cámara CCD SONY XC75CE con una óptica de 7,5mm y una tarjeta de adquisición de imágenes Matrox Corona alojada en el PC de Control.

En el PC de control también se aloja una tarjeta de entrada/salida (Analógico/Digital) modelo CIO-DAS 1600 de la empresa ComputerBoards utilizada para suministrar tensión al coche perseguidor. La salida de la tarjeta está configurada en el rango unipolar de 0 a 10 voltios. El modelo utilizado ofrece una intensidad de corriente de salida máxima de 20 miliamperios lo que precisó del uso de un amplificador de corriente cuyo esquema se detalla en la figura 2.

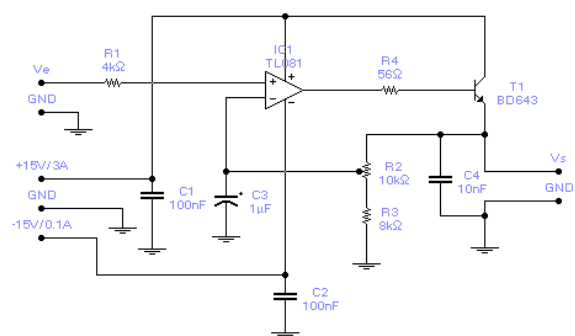


Figura 2: Esquema del amplificador

2.2 ESQUEMA DE BLOQUES DEL SISTEMA DE CONTROL VISUAL

En el diagrama de bloques de la figura 3 se representa el sistema controlado.

En este esquema se observa como la entrada al sistema es la señal D_{ref} que indica la distancia que se desea que mantengan los dos coches en movimiento. El error del sistema, e , se define como la diferencia entre D_{ref} y la distancia actual entre los coches. Este error es la entrada a un regulador $R(s)$ que tiene como salida una tensión V_1 que se ejercerá sobre el motor del coche perseguidor.

En la realimentación del sistema se dispone de un sistema de visión artificial para determinar la posición de los coches en movimiento, este procesamiento de imágenes se ha simbolizado en el

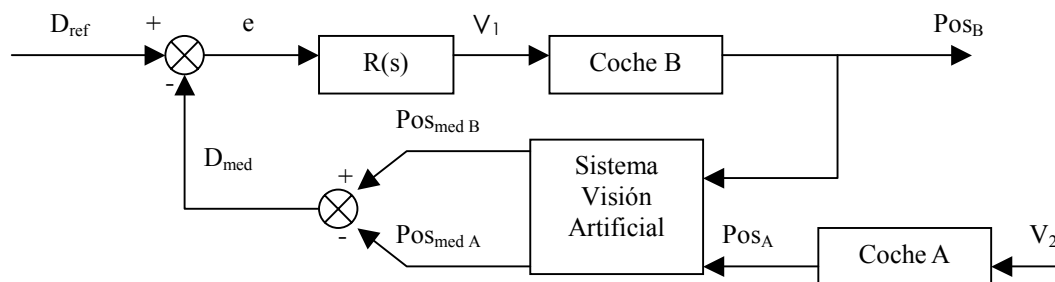


Figura 3: Diagrama de bloques del sistema controlado por posición

esquema 3 con el bloque “Sistema de Visión Artificial” y contiene los módulos de alto y bajo nivel encargados de extraer las características del movimiento.

A continuación se detallan los módulos de tracking encargados de las tareas de posicionamiento de los coches en el bucle de realimentación a través de técnicas de procesamiento de imágenes.

También se describirá el módulo de control encargado de gobernar de forma automática el coche perseguidor. En este apartado se analizará el desempeño del sistema para distintos reguladores $R(s)$ diseñados, indicando cual de ellos es el más adecuado.

3 MÓDULO DE TRACKING

En este módulo se realizan todas las tareas de visión artificial. Estas tareas incluyen procesamientos, tanto de bajo como de alto nivel, de las secuencias de imágenes conducentes a la obtención de la posición de los coches, y por ende a la distancia existente entre ellos.

3.1 MÓDULOS DE BAJO NIVEL

Dentro de este grupo se encuentran los módulos encargados de realizar el tracking, es decir, distinguir los objetos en movimiento a lo largo de la secuencia de imágenes. El algoritmo se basa en considerar que en el seguimiento los únicos objetos que van a sufrir algún movimiento son los que precisamente están implicados en el seguimiento, es decir, el coche perseguidor y el perseguido. Con esta consideración, si eliminamos la información de los elementos de cada imagen que no varían de una a otra tendremos únicamente los elementos móviles.

La primera fase del algoritmo básicamente consiste en establecer una metodología para aislar el movimiento de lo que permanece inmóvil durante todo el seguimiento, a través de procesamiento de imágenes a bajo nivel. La estrategia de procesamiento de las imágenes debe ser suficientemente rápida para que los retardos no sean elevados y a la vez sea robusta para la aplicación.

En la secuencia de imágenes de la figura 4 se observa el proceso que se lleva a cabo para determinar los centros de gravedad de los coches. Terminado este proceso se obtienen dos zonas compactas cuyo centro de gravedad corresponde con la posición de los coches. En el paso 4.E, para obtener una zona compacta correspondiente al objeto en movimiento, se dilata la imagen binaria utilizando un elemento estructurante circular.

En el transcurso del procesamiento de secuencias de imágenes puede aparecer ruido (u oclusiones) que provoque una incorrecta detección de movimiento. En estos casos, los módulos de nivel superior serán los encargados de analizar los centros obtenidos (algunos de ellos incorrectos debido al ruido) y eliminar los que en ningún caso se puedan tratar de movimiento de los coches.

3.2 MÓDULOS DE ALTO NIVEL

En este apartado se incluyen aquellos módulos que se podrían encuadrar dentro de la visión de alto nivel. De los módulos de bajo nivel se dispone de la información relativa a los centros de gravedad de los objetos móviles. Sin embargo, esta información puede no ser del todo correcta y hay que procesarla con los siguientes fines:

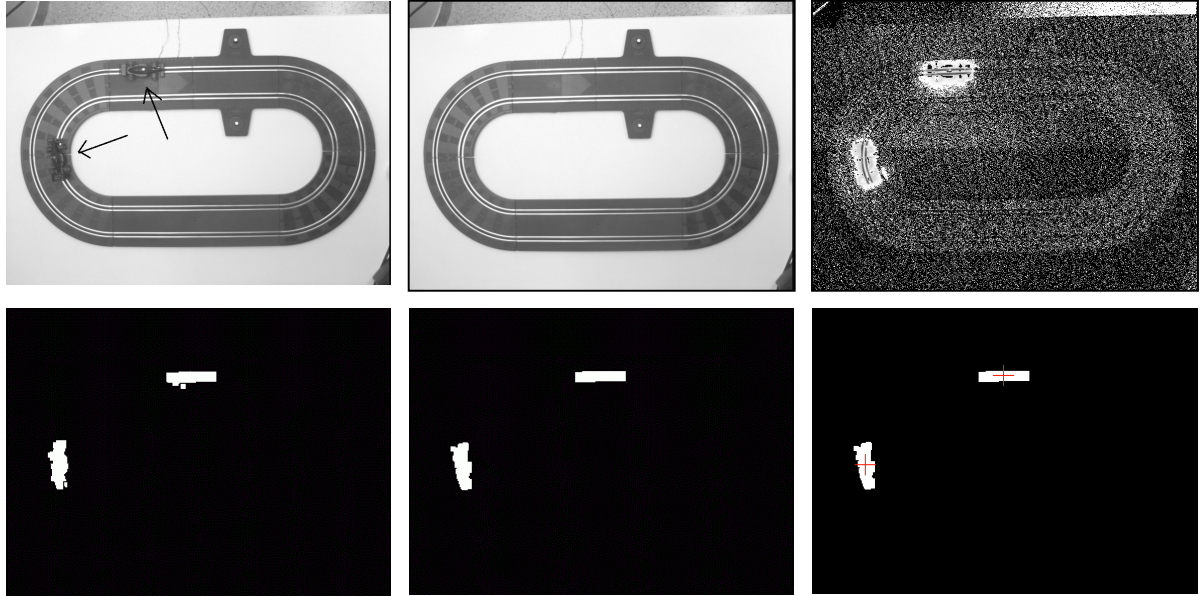


Figura 4: 4.A (Sup-Izq) Imagen con los coches en movimiento. 4.B (Sup-Cent) Imagen de fondo. 4.C (Sup-Der) Imagen resultado de diferencia en valor absoluto. 4.D (Inf-Izq) Imagen umbralizada. 4.E (Inf-Cent) Imagen diferencia umbralizada y dilatada. 4.F (Inf-Der) Centros de gravedad.

1. Si se ha producido algún defecto durante el proceso de captura o algún otro ruido, los módulos de nivel inferior no podrán distinguir a ciencia cierta cual de los centros de gravedad obtenidos son los realmente correctos. Por lo tanto, llegados a este punto, se deberá realizar una discriminación de la información imprecisa o incorrecta y aislar la que realmente interesa para el seguimiento.

Para resolver este problema, es necesario discriminar la información incorrecta para quedarse únicamente con los centros de masas de los coches. Para realizar esta tarea se ha seguido un modelo basado en predicción de movimiento, introduciendo una variable de ruido gaussiano cuya varianza se determinará empíricamente, introduciendo suficiente variedad en la predicción para captar posibles errores en la predicción.

El proceso de discriminación de centros de gravedad incorrectos parte de que se conoce perfectamente cual fue la posición de los coches en instantes anteriores. Visto esto, se realiza la suposición de que los coches se moverán en el instante actual con la misma velocidad que llevaron anteriormente más un cierto ruido.

Para predecir donde se va a encontrar el centroide de cada coche en el instante $t+2$ se realiza un proceso basado en que se conoce donde estuvo en los instantes anteriores (t y $t+1$).

Si consideramos a V como un vector de movimiento, x' e y' las coordenadas de un coche en el instante $t+2$, x e y las coordenadas en el instante $t+1$. Se tiene lo siguiente:

$$\begin{aligned} x' &= x + V_x \\ y' &= y + V_y \end{aligned} \quad (1)$$

Como las coordenadas (x, y) son conocidas porque son de un instante anterior, falta describir el cálculo del vector de movimiento $V (V_x, V_y)$.

Se va a considerar en la predicción que el movimiento que describe el móvil es el mismo que describía anteriormente más un cierto ruido. Para implementar esto tenemos lo siguiente:

$$\begin{aligned} V_x &= M_x + S_x \\ V_y &= M_y + S_y \end{aligned} \quad (2)$$

Siendo M_x y M_y el movimiento descrito por el coche entre los frames en t y $t+1$. A este movimiento se le suma un cierto ruido al que se ha denominado S_x y S_y . Por lo tanto, si se considerara como modelo de movimiento M_x y M_y , se estaría considerando que el móvil describirá en $t+2$ un movimiento exactamente igual al que describió para ir de t a $t+1$. Sin embargo, se le suma un factor de ruido (3) que se distribuye normalmente $N(0,1)$ para captar posibles modificaciones en el movimiento.

$$\begin{aligned} S_x &= N(0,1)*dx \\ S_y &= N(0,1)*dy \end{aligned} \quad (3)$$

,siendo dx y dy la desviación típica del movimiento. Esta desviación típica ha sido calculada empíricamente observando las desviaciones de velocidad entre dos instantes de tiempo que ocurren a lo largo de un seguimiento. $N(0,1)$ es un valor aleatorio de una normal con media 0 y desviación típica 1.

De esta manera se van prediciendo dónde van a estar los centroides de cada coche a partir de los centroides calculados entre los dos frames de instantes anteriores. Una vez realizada la predicción, la manera de discriminar los centroides verdaderos de los falsos consiste en que los verdaderos serán los que tengan las coordenadas más próximas a las que se predijeron. Por lo tanto, terminado este paso se puede garantizar que únicamente tenemos dos centros de masas y se corresponden con los de los coches.

Los siguientes pasos vienen encaminados a determinar el desplazamiento que puede haber sufrido la pista respecto a una posición de referencia así como la determinación de la distancia entre los dos coches.

2. Para dotar de la mayor flexibilidad posible al programa se ha considerado la posibilidad de que se produzcan desplazamientos y giros de la pista. Este hecho no es considerado en los módulos de nivel inferior, y será ahora cuando, realizando las operaciones oportunas, se corrijan posibles desplazamientos y giros.

Para determinar si se ha producido un desplazamiento en la pista se dispone de dos marcas pintadas de blanco en la misma, de manera que, detectando estas marcas se podrá establecer si la pista ha sufrido un desplazamiento y/o una rotación.

Para detectar la presencia de las marcas se procede a un proceso de umbralización que se representa esquemáticamente en la secuencia de imágenes 5.

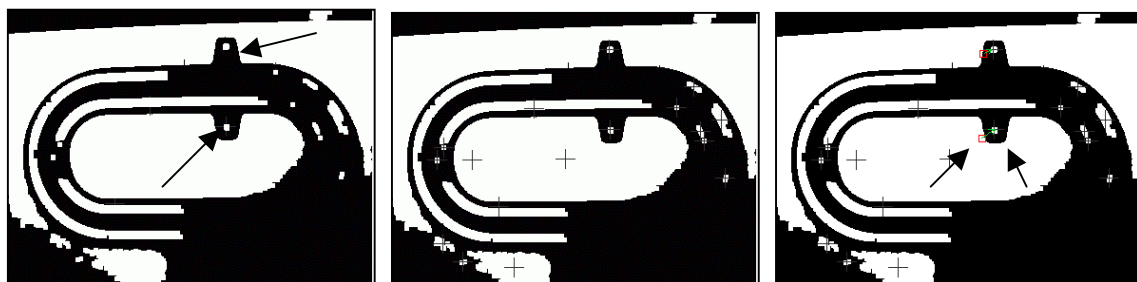


Figura 5. 5.A(izq) Imagen umbralizada. 5.B. (Cent) Imagen tras la detección de los centros. 5.C (Der) distancia entre la posición teórica de las marcas y la posición real.

Determinando la diferencia entre la posición teórica de las marcas y la que realmente se obtiene, se puede determinar qué desplazamiento ha sufrido la pista respecto a una posición de referencia. En general habrá sufrido una transformación simbolizada por una matriz $H = [R, T]$. Donde R representa la rotación y T la traslación.

3. Una vez que se dispone de manera fiable de los centros de gravedad de los móviles, el siguiente paso es el obtener la distancia entre ambos. Este paso no es inmediato ya que la distancia que separa los dos coches no es, casi siempre, la línea recta. Es necesario establecer algún criterio para calcular lo más rápidamente posible esta distancia.

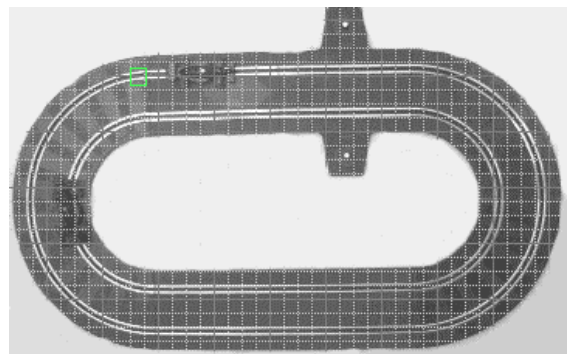


Figura 6: Pista calibrada con la cuadrícula. Se muestra la posición de referencia

Para medir la distancia entre los coches se ha realizado un proceso previo de calibración de la pista. Esta calibración consiste en definir una especie de cuadrícula que engloba los carriles de la pista. Con esto se dispone de una tabla que, para cada posición de la cuadrícula definida, da una distancia respecto de una posición de referencia.

Una vez determinados los centros de gravedad de los coches (P_i), consultando la posición ($H \cdot P_i$) en la tabla de distancias se podrá conocer la distancia de cada coche a la posición de referencia. Realizando

la diferencia entre ambos valores ya se dispondrá de la distancia real entre ambos coches.

4 MÓDULO DE CONTROL

Este módulo es el encargado de controlar el coche perseguidor según unas especificaciones dadas.

Se ha identificado, a través de técnicas en el dominio temporal, que la función de transferencia en bucle abierto del sistema representado en la figura 3 es la que se muestra en la ecuación (4):

$$G(s) = \frac{1.025 - 25.625s}{(1 + 25s)(1 + 67s)(1 + 98.53s)(1 + 74.64s)} \quad (4)$$

El sistema de visión artificial se ha modelado como un retardo ($T = 50$ mseg):

$$G(s) = e^{-Ts} \quad (5)$$

La función de transferencia de los motores de los coches eléctricos se ha determinado de forma empírica con técnicas en el dominio del tiempo.

Para la regulación del sistema se va a considerar $t_s = 500$ mseg y $M_p = 5\%$, así como un error en régimen permanente de posición E_p , del 10%. Los distintos reguladores diseñados son los siguientes:

$$R_{PD}(s) = 0.25 \frac{s + 0.01015}{s + 0.005329} \quad (6)$$

$$R_{PI}(s) = \frac{s + 0.0010472}{s + 0.0000562} \quad (7)$$

$$R_{PID}(s) = \frac{s + 0.01015}{s + 0.005329} \cdot \frac{s + 0.0010472}{s + 0.0000562} \quad (8)$$

En la figura 7 se muestra la respuesta ante un escalón de tensión para el sistema original y el sistema compensado con los reguladores PD (6) y PID (8).

Para el control real del sistema se han convertido las funciones de transferencia anteriores al dominio Z según [4]:

$$s = \frac{2 \cdot (1 - z^{-1})}{T_s (1 + z^{-1})} \quad (9)$$

5 RESULTADOS

La utilización de un regulador proporcional en el seguimiento implica un transitorio muy oscilatorio lo que provoca que el coche perseguidor haga continuas aceleraciones y frenadas para adaptarse a la velocidad del coche perseguido. Vistas las excesivas oscilaciones obtenidas durante la práctica se pasó a la implementación del regulador PD para intentar aumentar la velocidad del sistema ciñéndose a la vez a unas especificaciones de régimen transitorio que garantizara no obtener la sobreoscilación presente en una regulación de tipo proporcional

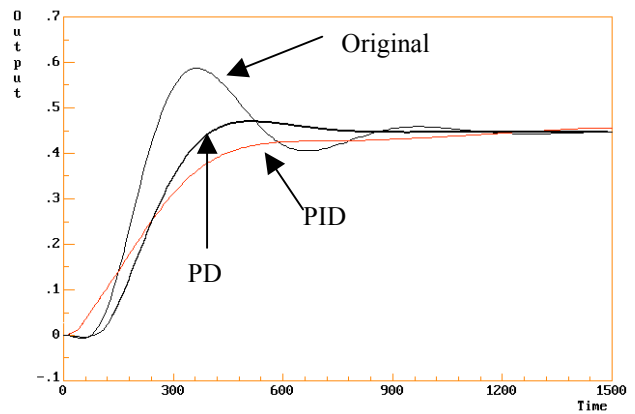


Figura 7: Comparativa de la respuesta ante escalón de tensión entre el sistema original y el compensado con los reguladores PD y PID diseñados

La implementación directa del PD calculado, proporciona mejores resultados que el regulador proporcional. Esta mejora se hace más patente cuanto menor es la distancia entre ambos coches, es decir, en la aproximación al vehículo perseguido. Utilizando el regulador PD se obtienen aceleraciones suficientemente rápidas a la vez que suaves para conseguir que se cumplan las especificaciones de distancia entre los coches durante el seguimiento.

En los resultados prácticos se ha comprobado que con la utilización de un regulador de tipo PID no se obtiene una mejora ostensible respecto a la utilización de una regulación diferencial. Esto es debido a que la utilización de la regulación integral es usada para mejorar las especificaciones de régimen permanente. Los continuos rozamientos, inercias, y cambios de velocidad, provocan que una modificación en el régimen permanente no sea tan clara y visible en la práctica.

En cuanto al algoritmo de tracking empleado se ha observado que los módulos de nivel inferior son muy sensibles al ruido, sin embargo, el posterior

procesamiento por los módulos de nivel superior eliminan en gran parte estos errores, así como que hace posible el seguimiento ante pequeñas variaciones de la iluminación. También es capaz de realizar el seguimiento ante pequeñas oclusiones (de corta duración).

Es de destacar que en la implementación actual se obtienen velocidades de procesamiento en torno a los 50 milisegundos por frame, velocidad de procesamiento suficientemente elevada para realizar el control en tiempo real.

6 CONCLUSIONES

La necesidad del procesamiento de imágenes a alta velocidad en los sistemas de control visual, provoca que el método seleccionado para la determinación de las características buscadas deba ser suficientemente simple y flexible para que no se produzcan excesivos retardos en el procesamiento de las imágenes y a la vez se capte la información buscada en presencia de ruidos.

Con el método propuesto de tracking, se ha podido realizar el control visual, así como la comparación en el funcionamiento del sistema utilizando distintos reguladores. La utilización de dos niveles en el procesamiento de las imágenes obliga a que los módulos de nivel superior utilicen conocimiento del entorno para aislar errores introducidos por los módulos de nivel inferior. Estos niveles superiores son capaces de, a la vez de eliminar ruidos, realizar el seguimiento salvando posibles oclusiones.

En la actualidad se está trabajando en la mejora de los niveles inferiores para hacerlos menos sensibles a variaciones en el entorno como pueden ser defectos en la iluminación.

Referencias

- [1] C. Bregler, J. Malik "Tracking people with twists and exponential maps" Proc IEEE Computer Society Conf. Computer Vision Pattern Recognition, 1998.
- [2] P. Corke, J. Roberts, G. Winstanley, "Vision-Based Control for Mining Automation" IEEE Robot. Automat. Vol 5, no. 4, pp. 44 - 49. Diciembre 1998.
- [3] Peter I Corke y Malcolm C. Good, "Dynamic Effects in Visual Closed-Loop Systems". Proc IEEE Transactions on robotics and automation Vol 12, NO 5, Octubre 1996
- [4] Ken Dutton, Steve Thompson, Bill Barraclough. "The Art of Control Engineering". Addison-Wesley, 1997
- [5] J. T. Feddema y R. W. Simon, "Visual Servoing and CAD-Driven Microassembly." IEEE . Robot Automat. Vol 5, no. 4, pp. 18-24 Diciembre 1998..
- [6] J.Hill y W. T. Park. "Real time control of a robot with a mobile camera", in Proc 9th ISIR, Washintong DC Mar 1979 pp, 233-246
- [7] Isard, M., Blake, A. CONDENSATION: Conditional Density Propagation for Visual Tracking. Int. Journal of Computer Vision. 29(1): 5 – 28, 1998.
- [8] K. Nickels, S. Hutchinson, "Model-Based Tracking of Complex Articulated Objects" IEEE Transaction on Robotics And Automation Vol. 17 no 1, Febrero 2001. Pp 28 – 36
- [9] J. Pomares. "Seguimiento de objetos en movimiento usando realimentación visual" Sistemas Informáticos. Dpto. Física, Ingeniería de Sistemas y Teoría de la Señal (Universidad de Alicante) 1999.
- [10] Chistopher Rasmussen. "Visual servoing & Mobile Robot Navigation" Computer Science 690 Junio 15, 1995.